

LSV models with stochastic interest rates and correlated jumps

Andrey Itkin

New York University, Tandon School of Engineering

December 15, 2015

Abstract

Pricing and hedging exotic options using local stochastic volatility models drew a serious attention within the last decade, and nowadays became almost a standard approach to this problem. In this paper we show how this framework could be extended by adding to the model stochastic interest rates and correlated jumps in all three components. We also propose a new fully implicit modification of the popular Hundsdorfer and Verwer and Modified Craig-Sneyd finite-difference schemes which provides second order approximation in space and time, is unconditionally stable and preserves positivity of the solution, while still has a linear complexity in the number of grid nodes.

Pricing and hedging exotic options using local stochastic volatility (LSV) models drew a serious attention within the last decade, and nowadays became almost a standard approach to this problem. For the detailed introduction into the LSV among multiple available references we mention a recent comprehensive literature overview in [Homescu, 2014]. Note, that the same model or its flavors appear in the literature under different names, such as stochastic local volatility model, universal volatility model of [Lipton, 2002], unspanned stochastic local volatility model (USLV) of [Halperin and Itkin, 2013], etc.

Despite LSV has a lot of attractive features allowing simultaneous pricing and calibration of both vanilla and exotic options, it was observed that in many situations, e.g., for short maturities, jumps in both the spot price and the instantaneous variance need to be taken into account to get a better replication of the market data on equity or FX derivatives. This approach was pioneered by [Bates, 1996] who extended the Heston model by introducing jumps with finite activity into the spot price (a jump-diffusion model). Then [Lipton, 2002] further extended this approach by considering local stochastic volatility to be incorporated into the jump-diffusion model (for the extension to an arbitrary Lévy model, see, e.g., [Pagliarani and Pascucci, 2012]). Later [Sepp, 2011b, Sepp, 2011a] investigated exponential and discrete jumps in both the underlying spot price S and the instantaneous variance v , and concluded that infrequent negative jumps in the latter are necessary to fit

the market data on equity options¹. In [Durhama and Park, 2013] a similar approach was proposed to use general jump-diffusion equations for modeling both S and v .

Note, that in the literature jump-diffusion models for both S and v are also known under the name SVCJ (stochastic volatility with contemporaneous jumps). These models as applied to pricing American options were intensively studied in [Salmi et al., 2014], for basket options in [Shirava and Takahashi, 2013].

Another way to extend the LSV model is to assume that the short interest rates r could be stochastic. Under this approach jumps are ignored, but instead a system of three Stochastic Differential Equations (SDE) with drifts and correlated diffusions is considered, see [Giese, 2006, Medvedev and Scaillet, 2010, Grzelak and Oosterlee, 2011, Haentjens and In't Hout, 2012, Hilpisch, 2011, Chiarella and Kang, 2013, Boyarchenko and Levendorskii, 2013] and references therein.

As we have already mentioned, accounting for jumps could be important to calibrate the LSV model to the market data. And making the interest rate stochastic doesn't violate this conclusion. Moreover, jumps in the interest rate itself could be important. For instance, in [Chen and Scott, 2004] a stochastic volatility model with jumps in both rates and volatility was calibrated to the daily data for futures interest rates in four major currencies which provided a better fit for the empirical distributions. Also the results in [Johannes, 2004] obtained using Treasury bill rates find evidence for the presence of jumps which play an important statistical role. Also it was found that jumps generally have a minor impact on yields, but they are important for pricing interest rate options.

In FX world there exist some variations of the discussed models. For instance, in [Doffou and Hillard, 2001] foreign and domestic interest rates are stochastic with no jumps while the exchange rate is modeled by jump-diffusion. In [Carr and Wu, 2004] both domestic and foreign rates were represented as a Lévy process with the diffusion component using a time-change approach. The diffusion components could be correlated in contrast to the jump components.

In the bond market, as shown in [Das, 2002], the information surprises result in discontinuous interest rates. In that paper a class of Poisson–Gaussian models of the Fed Funds rate was developed to capture the surprise effects. It was shown that these models offer a good statistical description of a short rate behavior, and are useful in understanding many empirical phenomena. Jump (Poisson) processes capture empirical features of the data which would not be captured by Gaussian models. Also there is strong evidence that the existing Gaussian models would be well-enhanced by jump and ARCH-type processes.

Overall, it would be desirable to have a model where the LSV framework could be combined with stochastic rates and jumps in all three stochastic drivers. We also want to treat these jumps as general Lévy processes, so not limiting us by only the jump-diffusion models. In addition, we consider Brownian components to be correlated as well as the jumps in all stochastic drivers to be correlated, while the diffusion and jumps remain uncorrelated. Fi-

¹Here we don't discuss this conclusion. However, for the sake of reference note, that this could be dictated by some inflexibility of the Heston model where vol-of-vol is proportional to $v^{0.5}$. More flexible models which consider the vol-of-vol power to be parameter of calibration, [Gatheral, 2008, Itkin, 2013], might not need jumps in v . See also [Sepp, 2014] and the discussion therein.

nally, since such a model is hardly analytically tractable when parameters of the model are time-dependent (which is usually helpful to better calibrate the model to a set of instruments with different maturities, or to a term-structure of some instrument), we need an efficient numerical method for pricing and calibration.

For this purpose in this paper we propose to exploit our approach first elaborated on in [Itkin and Lipton, 2015] for modeling credit derivatives. In particular, in the former paper we considered a set of banks with mutual interbank liabilities whose assets are driven by correlated Lévy processes. For every asset, the jumps were represented as a weighted sum of the common and idiosyncratic parts. Both parts could be simulated by an arbitrary Lévy model which is an extension of the previous approaches where either the discrete or exponential jumps were considered, or a Lévy copula approach was utilized. We provided a novel efficient (linear complexity in each dimension) numerical (splitting) algorithm for solving the corresponding 2D and 3D jump-diffusion equations, and proved its convergence and second order of accuracy in both space and time. Test examples were given for the Kou model, while the approach is in no way limited by this model.

In this paper we demonstrate how a similar approach can be used together with the Metzler model introduced by [Schoutens and Teugels, 1998, Schoutens, 2001]. It is built based on the Meixner distribution which belongs to the class of the infinitely divisible distributions. Therefore, it gives rise to a Lévy process - the Meixner process. The Meixner process is flexible and analytically tractable, i.e. its pdf and CF are known in closed form (in more detail see, e.g., [Itkin, 2014b] and references therein). The Meixner model is known to be rich and capable to be calibrated to the market data. Again, this model is chosen only as an example, because, in general, the approach in use is rather universal.

We also propose a new fully implicit modification of the popular Hundsdorfer and Verwer and Modified Craig-Sneyd finite-difference schemes which provides second order approximation in space and time, is unconditionally stable and preserves positivity of the solution, while still has a linear complexity in the number of grid nodes. This modification allows elimination of first few Rannacher steps as this is usually done in the literature to provide a better stability (see survey, e.g., in [Haentjens and In't Hout, 2012]), and provides much better stability of the whole scheme which is important when solving multidimensional problems.

The rest of the paper is organized as follows. In the next section we describe the model. Section 2 consists of two subsections. The first one introduces the new splitting method, which treats mixed derivatives terms implicitly, thus providing a much better stability. The second subsection describes how to deal with jumps if one uses the Meixner model. However, by no means this approach is restricted just by this model as, e.g., in [Itkin and Lipton, 2015] we used the Kou jump models using the same treatment of the jump terms. So here the Meixner model is taken as another example. Section 3 presents the results of some numerical experiments where prices of European vanilla and barrier options were computed using these model and numerical method. The final section concludes.

1 Model

We consider an LSV model with stochastic interest rates and jumps by introducing stochastic dynamics for variables S_t, v_t, r_t . We assume that it could include both diffusion and jumps components, as follows:

$$\begin{aligned} dS_t &= (r_t - q)S_t dt + \sigma_s(S_t, t)S_t^c \sqrt{v_t} W_s + S_t dL_{S_t, t}, \\ dv_t &= \kappa_v(t)[\theta_v(t) - v_t]dt + \xi_v v_t^a W_v + v_t dL_{v_t, t}, \\ dr_t &= \kappa_r(t)(\theta_r(t) - r_t)dt + \xi_r r_t^b W_r + r_t dL_{r_t, t}. \end{aligned} \quad (1)$$

Here q is the continuous dividend, t is the time, σ_s is the local volatility function, W_s, W_v, W_r are correlated Brownian motions, such that $\langle dW_i, dW_j \rangle = \rho_{ij}dt$, $i, j \in [s, v, r]$, $\kappa_v, \theta_v, \xi_v$ are the mean-reversion rate, mean-reversion level and volatility of volatility (vol-of-vol) for the instantaneous variance v_t , $\kappa_r, \theta_r, \xi_r$ are the corresponding parameters for the stochastic interest rate r_t , $0 \leq a < 2$, $0 \leq b < 2$, $0 \leq c < 2$ are some power constants which are introduced to add additional flexibility to the model as compared with the popular Heston ($\alpha = 0.5$), lognormal ($\alpha = 1$) and 3/2 ($\alpha = 1.5$) models². Processes L_s, L_v, L_r are pure discontinuous jump processes with generator A

$$Af(x) = \int_{\mathbb{R}} (f(x+y) - f(x) - y\mathbf{1}_{|y|<1}) \mu(dy),$$

with $\mu(dy)$ be a Lévy measure, and

$$\int_{|y|>1} e^y \mu(dy) < \infty.$$

At this stage, the jump measure $\mu(dx)$ is left unspecified, so all types of jumps including those with finite and infinite variation, and finite and infinite activity could be considered here.

Following [Itkin and Lipton, 2015] we introduce correlations between all jumps as this was done in [Ballotta and Bonfiglioli, 2014]. They construct the jump process as a linear combination of two independent Lévy processes representing the systematic factor and the idiosyncratic shock, respectively (see also [Cont and Tankov, 2004]). It has an intuitive economic interpretation and retains nice tractability, as the multivariate characteristic function in this model is available in closed form based on the following proposition of [Ballotta and Bonfiglioli, 2014]:

Proposition 1.1 *Let $Z_t, Y_{j,t}$, $j = 1, \dots, n$ be independent Lévy processes on a probability space (Q, F, P) , with characteristic functions $\phi_Z(u; t)$ and $\phi_{Y_j}(u; t)$, for $j = 1, \dots, n$ respectively. Then, for $b_j \in \mathbb{R}$, $j = 1, \dots, n$*

$$X_t = (X_{1,t}, \dots, X_{n,t})^\top = (Y_{1,t} + b_1 Z_t, \dots, Y_{n,t} + b_n Z_t)^\top$$

²If, however, somebody wants to determine these parameters by calibration, she has to be careful, because having both vol-of-vol and a power constant in the same diffusion term brings an ambiguity into the calibration procedure. Nevertheless, this ambiguity can be resolved if for calibration some additional financial instruments are used, e.g., exotic option prices are combined with the variance swaps prices, see [Itkin, 2013].

is a Lévy process on \mathbb{R}^n . The resulting characteristic function is

$$\phi_{\mathbf{X}}(\mathbf{u}; t) = \phi_Z \left(\sum_{i=1}^n b_i u_i; t \right) \prod_{i=1}^n \phi_{Y_j}(u_j; t), \quad \mathbf{u} \in \mathbb{R}^n.$$

By construction every factor $X_{i,t}, i = 1, \dots, n$ includes a common factor Z_t . Therefore, all components $X_{i,t}, i = 1, \dots, n$ could jump together, and loading factors b_i determine the magnitude (intensity) of the jump in $X_{i,t}$ due to the jump in Z_t . Thus, all components of the multivariate Lévy process \mathbf{X}_t are dependent, and their pairwise correlation is given by (again see [Ballotta and Bonfiglioli, 2014] and references therein)

$$\rho_{j,i} = \text{Corr}(X_{j,t}, X_{i,t}) = \frac{b_j b_i \text{Var}(Z_1)}{\sqrt{\text{Var}(X_{j,1})} \sqrt{\text{Var}(X_{i,1})}}.$$

Such a construction has multiple advantages, namely:

1. As $\text{sign}(\rho_{i,j}) = \text{sign}(b_i b_j)$, both positive and negative correlations can be accommodated
2. In the limiting case $b_i \rightarrow 0$ or $b_j \rightarrow 0$ or $\text{Var}(Z_1) = 0$ the margins become independent, and $\rho_{i,j} = 0$. The other limit $b_i \rightarrow \infty$ or $b_j \rightarrow \infty$ represents a full positive correlation case, so $\rho_{i,j} = 1$. Accordingly, $b_i \rightarrow \infty, b_{3-i} \rightarrow \infty, i = 1, 2$ represents a full negative correlation case as in this limit $\rho_{i,j} = -1$.

According to this setup, the total instantaneous correlation between the assets x_i and x_j reads [Ballotta and Bonfiglioli, 2014]

$$\tilde{\rho}_{ij} = \frac{\rho \sigma_i \sigma_j + b_i b_j \text{Var}(Z_1)}{\sqrt{\sigma_i^2 + \text{Var}(X_{i,1})} \sqrt{\sigma_j^2 + \text{Var}(X_{j,1})}}. \quad (2)$$

To price contingent claims, e.g., vanilla or exotic options written on the underlying spot price, by using a standard technique as in [Cont and Tankov, 2004], the following multidimensional PIDE could be derived which describes the evolution of the option price V under risk-neutral measure

$$V_\tau = [\mathcal{D} + \mathcal{J}]V, \quad (3)$$

where $\tau = T - t$ is the backward time, T is the time to the contract expiration, \mathcal{D} is the

three-dimensional linear convection-diffusion operator of the form

$$\begin{aligned}
\mathcal{D} &= F_0 + F_1 + F_2 + F_3, \\
F_1 &= (r - q)S \frac{\partial}{\partial S} + \frac{1}{2}\sigma_s^2 S^{2c} v \frac{\partial^2}{\partial S^2}, \\
F_2 &= \kappa_v(t)[\theta_v(t) - v] \frac{\partial}{\partial v} + \frac{1}{2}\xi_v^2 v^{2a} \frac{\partial^2}{\partial v^2}, \\
F_3 &= \kappa_r(t)[\theta_r(t) - r] \frac{\partial}{\partial r} + \frac{1}{2}\xi_r^2 r^{2b} \frac{\partial^2}{\partial r^2}, \\
F_0 &= \rho_{s,v}\sigma_s(S, t)\xi_v(t)S^c v^{a+0.5} \frac{\partial^2}{\partial S \partial v} + \rho_{s,r}\sigma_s(S, t)\xi_r(t)S^c \sqrt{v} r^b \frac{\partial^2}{\partial S \partial r} \\
&\quad + \rho_{v,r}\xi_v(t)\xi_r(t)v^a r^b \frac{\partial^2}{\partial r \partial v},
\end{aligned} \tag{4}$$

and \mathcal{J} is the jump operator

$$\begin{aligned}
\mathcal{J}V &= \int_{-\infty}^{\infty} \left[V(x_s + y_s, x_v + y_v, x_r + y_r, \tau) - V(x_s, x_v, x_r, \tau) \right. \\
&\quad \left. - \sum_{\chi \in [s, v, r]} (e^{y_\chi} - 1) \frac{\partial V(x_s, x_v, x_r, \tau)}{\partial \chi} \right] \mu(dy_s dy_v dy_r),
\end{aligned} \tag{5}$$

where $\mu(dy_s dy_v dy_r)$ is the three-dimensional Lévy measure, and $x_s = \log S/S_0$, $x_v = \log v/v_0$, $x_r = \log r/r_0$.

This PIDE has to be solved subject to the boundary and terminal conditions. We assume that the terminal condition for equity derivatives reads

$$V(S, v, r, T) = P(S),$$

where $P(S)$ is the option payoff as defined by the corresponding contract. The boundary conditions could be set, e.g., as in [Haentjens and In't Hout, 2012]. However, in the presence of jumps these conditions should be extended as follows. Suppose we want to use finite-difference method to solve the above PIDE and construct a jump grid, which is a superset of the finite-difference grid used to solve the diffusion equation (i.e. when $\mathcal{J} = 0$, see [Itkin, 2014a]). Then these boundary conditions should be set on this jump grid as well as at the boundaries of the diffusion domain.

2 Solution of the PIDE

To solve Eq.(3) we use a splitting algorithm described in [Itkin and Lipton, 2015]. The algorithm provides the second order approximation in time (assuming that at every splitting

step the corresponding problem is solved with the same order of approximation) and reads

$$\begin{aligned}
V(\tau + \Delta\tau) &= e^{0.5\Delta\tau\mathcal{D}} e^{0.5\Delta\tau\mathcal{J}_s} e^{0.5\Delta\tau\mathcal{J}_v} e^{0.5\Delta\tau\mathcal{J}_r} e^{\Delta\tau\mathcal{J}_{123}} \\
&\quad \cdot e^{0.5\Delta\tau\mathcal{J}_r} e^{0.5\Delta\tau\mathcal{J}_v} e^{0.5\Delta\tau\mathcal{J}_s} e^{0.5\Delta\tau\mathcal{D}} V(\tau), \\
J_\chi &= \phi_\chi(-i\nabla_\chi), \quad J_{123} = \phi_Z(-i \sum_{\chi \in [s,v,r]} b_\chi \nabla_\chi), \quad \nabla_\chi \equiv \frac{\partial}{\partial \chi}.
\end{aligned} \tag{6}$$

Thus, this requires a sequential solution of 9 equations at every time step. The first and the last steps are pure advection-diffusion problems and could be solved using, e.g., a finite difference method proposed in [Haentjens and In't Hout, 2012]. We, however, slightly modified it by replacing an explicit scheme for the mixed derivative operators with the implicit ones. The detailed description of this approach as well as our reasons for doing that are given in the next section.

2.1 Advection-diffusion problem

We follow [In't Hout and Welfert, 2007], who consider the unconditional stability of second-order finite-difference schemes used to numerically solve multi-dimensional diffusion problems containing mixed spatial derivatives. They investigate the ADI scheme proposed by Craig and Sneyd (see references in the paper), a modified version of Craig and Sneyd's ADI scheme, and the ADI scheme introduced by Hundsdorfer and Verwer. The necessary conditions are derived on the parameters of each of these schemes for unconditional stability in the presence of mixed derivative terms.

For example, let us choose a HV scheme. The main result of [In't Hout and Welfert, 2007] is that under some mild conditions on the parameter θ of the scheme, the second-order Hundsdorfer and Verwer (HV) scheme is unconditionally stable when applied to semi-discretized diffusion problems with mixed derivative terms in an arbitrary spatial dimension $k > 2$. For the 3D convection-diffusion problem in Eq.(3) with $\mathcal{J} = 0$, the HV scheme defines an approximation $V_n \approx V(\tau_n)$, $n = 1, 2, 3, \dots$, by performing a series of (fractional) steps:

$$\begin{aligned}
Y_0 &= V_{n-1} + \Delta\tau F(\tau_{n-1})V_{n-1}, \\
Y_j &= Y_{j-1} + \theta\Delta\tau [F_j(\tau_n)Y_j - F_j(\tau_{n-1})V_{n-1}], \quad j = 1, 2, 3 \dots, k, \\
\tilde{Y}_0 &= Y_0 + \frac{1}{2}\Delta\tau [F(\tau_n)Y_k - F(\tau_{n-1})V_{n-1}], \\
\tilde{Y}_j &= \tilde{Y}_{j-1} + \theta\Delta\tau [F_j(\tau_n)\tilde{Y}_j - F_j(\tau_n)Y_k], \quad j = 1, 2, 3 \dots, k, \\
V_n &= \tilde{Y}_k,
\end{aligned} \tag{7}$$

where $F = \sum_j F_j$, $j = 0, 1 \dots k$. This scheme is of order two in time for any value of θ , so this parameter can be chosen to meet additional requirements, [In't Hout and Welfert, 2007]. An advantage of this scheme is that the fractional steps with mixed derivatives, i.e. 1 and 5, are solved by using an explicit scheme. At the same time this could bring a problem, because a

very careful approximation of the mixed derivative term is required to preserve the stability and positivity of the solution³. Sometimes this requires very small step in time to be chosen.

In the 2D case to resolve this rather delicate issue a seven-point stencil for discretization of the mixed-derivative operator that preserves the positivity of the solution was proposed in [Toivanen, 2010, Chiarella et al., 2008] for correlations $\rho < 0$, and in [Ikonen and Toivanen, 2008, Ikonen and Toivanen, 2007] for positive correlations. However, in their schemes the mixed derivative term was treated implicitly (that is the reason they needed a discretized matrix to be an M-matrix). In our case the entire matrix in the right-hand side of steps 3,5 should be either a positive matrix, or a Metzler matrix (in this case the negative of an M-matrix). The latter can be achieved when using approximations of [Toivanen, 2010, Chiarella et al., 2008] and [Ikonen and Toivanen, 2008, Ikonen and Toivanen, 2007] in an opposite order, i.e. use approximations recommended for $\rho > 0$ when $\rho < 0$, and vice versa. However, due to the nature of the 7-point stencil, they are not able to provide a rigorous second order approximation of the mixed derivatives.

In our numerical experiments even using these explicit analogs of the mixed derivatives approximation in the 3D case was not always sufficient. Indeed, either we use real second order approximation of the mixed derivatives relying on the fact that in the HV splitting scheme F_0 comes only as a part of F . Hence, the negative terms in F_0 can be partly or even fully compensated by the other terms. Unfortunately, at some values of the model parameters this could be insufficient to provide the total positivity of the solution. Or we use the 7-points stencil which works well for the implicit scheme (for the reasons which will become clear in Appendix A when proving our Theorem), but still doesn't provide a necessary stability for the explicit scheme. Thus, one has to choose a very small step in time, which is impractical. Therefore, in this paper to provide an additional stability of the whole splitting scheme we modified this step as follows.

The main idea is to sacrifice the simplicity of the explicit representation of the mixed derivative term for the better stability. That is what was done in [Toivanen, 2010, Chiarella et al., 2008, Ikonen and Toivanen, 2008, Ikonen and Toivanen, 2007] who dealt with a 2D case and used an implicit approximation of the mixed derivatives term. However, in this paper we propose another approach.

Consider the first step in Eq.(7). Since here only the first order approximation in time is necessary, this step can be re-written in two steps

$$\begin{aligned} \frac{\partial V^*}{\partial \tau} &= F_0(\tau_{n-1}) \equiv F_{Sv}(\tau_{n-1}) + F_{Sr}(\tau_{n-1}) + F_{vr}(\tau_{n-1}), \\ V(\tau) &= V^*(\tau) + \Delta\tau[F_1(\tau_{n-1}) + F_2(\tau_{n-1}) + F_3(\tau_{n-1})]V^*(\tau), \end{aligned} \tag{8}$$

³This is especially important at the first few steps in time because of a step-function nature of the payoff. So a smoothing scheme, e.g., [Rannacher, 1984], is usually applied at the first steps, which, however, loses the second order approximation at these steps.

with

$$\begin{aligned}
F_{Sv}(\tau_{n-1}) &= \rho_{s,v} \sigma_s(S, \tau) S^c \xi_v(\tau) v^{a+0.5} \frac{\partial^2}{\partial S \partial v} \equiv \rho_{s,v} W(S) W(v) \frac{\partial^2}{\partial S \partial v}, \\
F_{Sr}(\tau_{n-1}) &= \rho_{s,r} \sigma_s(S, \tau) S^c \sqrt{v} \xi_r(\tau) r^b \frac{\partial^2}{\partial S \partial r} \equiv \rho_{s,r} W(S) W(r) \sqrt{v} \frac{\partial^2}{\partial S \partial r}, \\
F_{vr}(\tau_{n-1}) &= \rho_{v,r} \xi_v(\tau) v^a \xi_r(\tau) r^b \frac{\partial^2}{\partial r \partial v} \equiv \rho_{v,r} \frac{W(v) W(r)}{\sqrt{v}} \frac{\partial^2}{\partial r \partial v},
\end{aligned}$$

where $W(S) = \sigma(S, \tau) S^c$, $W(v) = \xi_v(\tau) v^{a+0.5}$, $W(r) = \xi_r(\tau) r^b$.

So efficiently at the first sub-splitting step we take a liberty to solve the first equation of Eq.(8) as we like, and the remaining part (the second sub-step) is treated explicitly, e.g. in the same way as in the HV scheme.

Now, a general solution of this first equation in Eq.(8) can be written in the operator form as

$$V(\tau + \Delta\tau) = e^{\Delta\tau(F_{Sv}(\tau_{n-1}) + F_{Sr}(\tau_{n-1}) + F_{vr}(\tau_{n-1}))} V(\tau).$$

Again, with $O(\Delta\tau)$

$$V(\tau + \Delta\tau) = e^{\Delta\tau F_{Sv}(\tau_{n-1})} e^{\Delta\tau F_{Sr}(\tau_{n-1})} e^{\Delta\tau F_{vr}(\tau_{n-1})} V(\tau),$$

or using splitting

$$\begin{aligned}
V^{(1)} &= e^{\Delta\tau F_{Sv}(\tau_{n-1})} V(\tau), \\
V^{(2)} &= e^{\Delta\tau F_{Sr}(\tau_{n-1})} V^{(1)}, \\
V(\tau + \Delta\tau) &= e^{\Delta\tau F_{vr}(\tau_{n-1})} V^{(2)}.
\end{aligned} \tag{9}$$

The order of the splitting steps usually doesn't matter.

Accordingly, it is sufficient to consider just one step in Eq.(9) since the others can be done in the similar way. For example, below let us consider step 1. First, we use Páde approximation (0,1) which provides approximation of the first line in Eq.(9) with the first order in $\Delta\tau$, and is implicit. Approximation wise this is equivalent to the first line of Eq.(7). Having that, the first equation in Eq.(9) transforms to

$$[1 - \Delta\tau \rho_{s,v} W(S) W(v) \nabla_S \nabla_v] V^{(1)} = V(\tau). \tag{10}$$

Second, we again rewrite it using a trick⁴

$$\begin{aligned}
&\left(P - \sqrt{\Delta\tau} \rho_{s,v} W(S) \nabla_S \right) \left(Q + \sqrt{\Delta\tau} W(v) \nabla_v \right) V^{(1)} \\
&= V(\tau) + \left[(PQ - 1) - Q \sqrt{\Delta\tau} \rho_{s,v} W(S) \nabla_S + P \sqrt{\Delta\tau} W(v) \nabla_v \right] V^{(1)},
\end{aligned} \tag{11}$$

⁴The trick is motivated by the desire to build an ADI scheme which consists of two one-dimensional steps, because for the 1D equations we know how to make the rhs matrix to be an EM-matrix [Itkin, 2014a].

where P, Q , are some positive numbers which have to be chosen based on some conditions, e.g., to provide diagonal dominance of the matrices in the parentheses in the lhs of Eq.(11), see below.

The intuition for this representation is as follows. Suppose we need to solve some parabolic PDE and represent the solution in the form of a matrix exponential $V(\tau + \Delta\tau) = e^{\Delta\tau J} V(\tau)$. Since computing the matrix exponential might be expensive, to preserve the second order approximation in $\Delta\tau$ one can use a second order Páde approximation. In this case, e.g., a popular Crank-Nicholson scheme preserves positivity of the solution only if the negative diagonal elements $d_{0,i}$, $i = 1, N$ of $\frac{1}{2}\Delta\tau J$ obey the condition $d_{0,i}(\frac{1}{2}\Delta\tau J) + 1 > 0, \forall i \in (1, N)$. This efficiently issues some limitations on the time step $\Delta\tau$. As a resolution, e.g., in [Wade et al., 2005] higher order fully implicit Páde approximations were proposed to be used instead of the Crank-Nicholson scheme. This solves the problem with getting a positive solution since

$$e^y \approx 1 - y + \frac{1}{2}y^2 + O(y^2) = \frac{1}{2}[y - (1 + \iota)][y - (1 - \iota)], \quad y \equiv \Delta\tau J, \quad \iota = \sqrt{-1},$$

and by using an appropriate discretization each matrix in the parentheses can be made an M-matrix which inverse is a non-negative matrix. This can be done when J is a 1D parabolic operator. Performance-wise, this, however, gives rise to solving few (e.g., 2 in the case of Páde (0,2) approximation) systems of linear equations with complex numbers. Hence, the complexity of the solution is, at least, 4 times worse. Our representation Eq.(11) aims to utilize a similar idea, but being transformed to the iterative method. The key point here is that we use a theory of EM-matrices (see [Itkin, 2014a, Itkin, 2015] and references therein), and manage to propose a second order approximation of the first derivative which makes our matrices to be real EM-matrices. So again, the inverse of the latter is a positive matrix.

Eq.(11) can be solved using fixed-point Picard iterations. One can start with setting $V(1) = V^0 = V(\tau)$ in the rhs of Eq.(11), then solve sequentially two systems of equations

$$\begin{aligned} (Q + \sqrt{\Delta\tau} W(v) \nabla_v) V^* &= V(\tau) + [PQ - 1 - Q\sqrt{\Delta\tau} \rho_{s,v} W(S) \nabla_S + P\sqrt{\Delta\tau} W(v) \nabla_v] V^k, \\ (P - \sqrt{\Delta\tau} \rho_{s,v} W(S) \nabla_S) V^{k+1} &= V^*. \end{aligned} \quad (12)$$

Here V^k is the value of $V^{(1)}$ at the k -th iteration.

Before constructing a finite difference scheme to solve this equation we need to introduce some definitions. Define a one-sided *forward* discretization of ∇ , which we denote as A_x^F : $A_x^F C(x) = [C(x + h, t) - C(x, t)]/h$. Also define a one-sided *backward* discretization of ∇ , denoted as A_x^B : $A_x^B C(x) = [C(x, t) - C(x - h, t)]/h$. These discretizations provide first order approximation in h , e.g., $\nabla C(x) = A_x^F C(x) + O(h)$. To provide the second order approximations, use the following definitions. Define $A_{2,x}^C = A_x^F A_x^B$ - the *central* difference approximation of the second derivative ∇_x^2 , and $A_x^C = (A_x^F + A_x^B)/2$ - the *central* difference approximation of the first derivative ∇ . Also define a one-sided second order approximations to the first derivatives: *backward* approximation $A_{2,x}^B$: $A_{2,x}^B C(x) = [3C(x) - 4C(x - h) + C(x - 2h)]/(2h)$, and *forward* approximation $A_{2,x}^F$: $A_{2,x}^F C(x) = [-3C(x) + 4C(x + h) -$

$C(x + 2h)]/(2h)$. Also I_x denotes a unit matrix. All these definitions assume that we work on a uniform grid, however this could be easily generalized for the non-uniform grid as well, see, e.g., [In't Hout and Foulon, 2010].

For solving Eq.(12) we propose two FD schemes. The first one (Scheme A) is introduced by the following Propositions⁵:

Proposition 2.1 *Let us assume $\rho_{s,v} \geq 0$, and approximate the lhs of Eq.(12) using the following finite-difference scheme:*

$$\begin{aligned} \left(QI_v + \sqrt{\Delta\tau}W(v)A_{2,v}^B \right) V^* &= \alpha^+ V(\tau) - V^k, \\ \left(PI_S - \sqrt{\Delta\tau}\rho_{s,v}W(S)A_{2,S}^F \right) V^{k+1} &= V^*, \\ \alpha^+ &= (PQ + 1)I_v - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_S^B + P\sqrt{\Delta\tau}W(v)A_v^F. \end{aligned} \quad (13)$$

Then this scheme is unconditionally stable in time step $\Delta\tau$, approximates Eq.(12) with $O(\sqrt{\Delta\tau} \max(h_S, h_v))$ and preserves positivity of the vector $V(x, \tau)$ if $Q = \beta\sqrt{\Delta\tau}/h_v$, $P = \beta\sqrt{\Delta\tau}/h_S$, where h_v, h_S are the grid space steps correspondingly in v and S directions, and the coefficient β must be chosen to obey the condition:

$$\beta > \max_{S,v} [W(v) + \rho_{s,v}W(S)].$$

Proof See Appendix A.

If $\sqrt{\Delta\tau} \approx \max(h_S, h_v)$, then the whole scheme becomes of the second order in space. However, this would be a serious restriction inherent to the explicit schemes. Therefore, in this paper we don't rely on it. Note, that in practice the time step is usually chosen such that $\sqrt{\Delta\tau} \ll 1$, and hence the whole scheme is expected to be closer to the second, rather than to the first order in h . As it has been already mentioned, this is similar to [Toivanen, 2010, Chiarella et al., 2008] for correlations $\rho < 0$, and [Ikonen and Toivanen, 2008, Ikonen and Toivanen, 2007] for positive correlations, where a seven point stencil breaks a rigorous second order of approximation in space.

A similar Proposition can be proved in case $\rho_{s,v} \leq 0$.

Proposition 2.2 *Let us assume $\rho_{s,v} \leq 0$, and approximate the lhs of Eq.(12) using the following finite-difference scheme of the second order in space:*

$$\begin{aligned} \left(QI_v + \sqrt{\Delta\tau}W(v)A_{2,v}^B \right) V^* &= \alpha^- V(\tau) - V^k, \\ \left(PI_S - \sqrt{\Delta\tau}\rho_{s,v}W(S)A_{2,S}^B \right) V^{k+1} &= V^*, \\ \alpha^- &= (PQ + 1)I_v - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_S^F + P\sqrt{\Delta\tau}W(v)A_v^F. \end{aligned} \quad (14)$$

⁵For the sake of clearness we formulate this Proposition for the uniform grid, but it should be pretty much transparent how to extend it for the non-uniform grid.

Then this scheme is unconditionally stable in time step $\Delta\tau$, approximates Eq.(12) with $O(\sqrt{\Delta\tau} \max(h_S, h_v))$ and preserves positivity of the vector $V(x, \tau)$ if $Q = \beta\sqrt{\Delta\tau}/h_v$, $P = \beta\sqrt{\Delta\tau}/h_S$, where h_v, h_S are the grid space steps correspondingly in v and S directions, and the coefficient β must be chosen to obey the condition:

$$\beta > \max_{S,v} [W(v) - \rho_{s,v} W(S)].$$

Proof The proof is completely analogous to that given for Proposition 2.1, therefore we omit it for the sake of brevity.

The results formulated in these two Propositions can be further improved by the cost of solving two additional systems of linear equations with a banded matrix, so efficiently by doubling the time necessary to solve Eq.(12). We call this FD scheme as Scheme B.

Proposition 2.3 Let us assume $\rho_{s,v} \geq 0$, and approximate the lhs of Eq.(12) using the following finite-difference scheme:

$$\begin{aligned} M^+ \left(QI_v + \sqrt{\Delta\tau} W(v) A_{2,v}^B \right) V^* &= V(\tau) - M^+ V^k, \\ \left(PI_S - \sqrt{\Delta\tau} \rho_{s,v} W(S) A_{2,S}^F \right) V^{k+1} &= V^*, \\ M^+ &= [\alpha^+]^{-1}, \quad \alpha^+ = (PQ + 1)I_v - Q\sqrt{\Delta\tau} \rho_{s,v} W(S) A_{2,S}^B + P\sqrt{\Delta\tau} W(v) A_{2,v}^F. \end{aligned} \tag{15}$$

Then this scheme is unconditionally stable in time step $\Delta\tau$, approximates Eq.(12) with $O(\max(h_S^2, h_v^2))$ and preserves positivity of the vector $V(x, \tau)$ if $Q = \beta\sqrt{\Delta\tau}/h_v$, $P = \beta\sqrt{\Delta\tau}/h_S$, where h_v, h_S are the grid space steps correspondingly in v and S directions, and the coefficient β must be chosen to obey the condition:

$$\beta > \max_{S,v} [W(v) + \rho_{s,v} W(S)].$$

The scheme Eq.(15) has a linear complexity in each direction.

Proof See Appendix B.

Proposition 2.4 Let us assume $\rho_{s,v} \leq 0$, and approximate the lhs of Eq.(12) using the following finite-difference scheme of the second order in space:

$$\begin{aligned} M^- \left(QI_v + \sqrt{\Delta\tau} W(v) A_{2,v}^B \right) V^* &= V(\tau) - M^- V^k, \\ \left(PI_S - \sqrt{\Delta\tau} \rho_{s,v} W(S) A_{2,S}^B \right) V^{k+1} &= V^*, \\ M^- &= [\alpha^-]^{-1}, \quad \alpha^- = (PQ + 1)I_v - Q\sqrt{\Delta\tau} \rho_{s,v} W(S) A_S^F + P\sqrt{\Delta\tau} W(v) A_v^F. \end{aligned} \tag{16}$$

Then this scheme is unconditionally stable in time step $\Delta\tau$, approximates Eq.(12) with $O(\sqrt{\Delta\tau} \max(h_S, h_v))$ and preserves positivity of the vector $V(x, \tau)$ if $Q = \beta\sqrt{\Delta\tau}/h_v$, $P =$

$\beta\sqrt{\Delta\tau}/h_S$, where h_v, h_S are the grid space steps correspondingly in v and S directions, and the coefficient β must be chosen to obey the condition:

$$\beta > \max_{S,v} [W(v) - \rho_{s,v} W(S)].$$

The scheme Eq.(16) has a linear complexity in each direction.

Proof The proof is completely analogous to that given for Proposition 2.3, therefore we omit it for the sake of brevity.

Once again we want to underline that the described approach to deal with the mixed derivative term supplies just the first order approximation in time. But that is exactly what was done in the HV scheme as well. Nevertheless the whole splitting scheme Eq.(7) is of the second order in $\Delta\tau$.

The coefficient β should be chosen experimentally. In our experiments described in the following sections we used

$$\beta = 10 \max_{S,v} [W(v) - \rho_{s,v} W(S)]. \quad (17)$$

For the second and third equations in Eq.(9) similar Propositions can be used to solve these equations and guarantee the second order approximation in space, the first order approximation in time and positivity of the solution as well as the convergence of the Picard fixed point iterations. A small but important improvement, however, must be made for the second equation in Eq.(9) since the definition of $F_{Sr}(\tau_{n-1})$ contains \sqrt{v} which is a dummy variable for this equation. Accordingly, as this equation should be solved in a loop on $v_j, j = 1, \dots, N_v$, where v_j are the nodes on the v -grid, and N_v is the number of these nodes, for each such a step its own β_j must be computed based on the condition

$$\beta_j > \max_{S,v} [W(v) - \rho_{s,v} W(S) \sqrt{v_j}].$$

This, however, doesn't bring any problem.

For the reference, we provide an explicit formulae for the first derivative for the backward D_{2B}^1 and forward D_{2F}^1 approximations of the second order at a non-uniform-grid. They read [Haentjens and In't Hout, 2012]

$$\begin{aligned} D_{2B}^1 f(x) \Big|_{x=x_i} &= f(x_i) \frac{h_i}{h_{i-1}(h_i + h_{i-1})} - f(x_{i-1}) \frac{h_{i-1} + h_i}{h_i h_{i-1}} + f(x_{i-2}) \frac{h_{i-1} + 2h_i}{h_i(h_{i-1} + h_i)}, \\ D_{2F}^1 f(x) \Big|_{x=x_i} &= -f(x_i) \frac{2h_{i+1} + h_{i+2}}{h_{i+1} h_{i+2}} + f(x_{i+1}) \frac{h_{i+2} + h_{i+1}}{h_{i+2} h_{i+1}} - f(x_{i+2}) \frac{h_{i+1}}{h_{i+2}(h_{i+2} + h_{i+1})}, \end{aligned}$$

where $h_i = f(x_i) - f(x_{i-1})$. Based on this definition, the matrices A_2^B, A_2^F can be constructed accordingly.

2.1.1 Fully implicit scheme

For even better stability, the whole first step Eq.(8) of the HV scheme can be made fully implicit. In doing that observe, that the first line in Eq.(7) is a Páde approximation (0,1) of the equation

$$\frac{\partial V(\tau)}{\partial \tau} = [F_0(\tau) + F_1(\tau) + F_2(\tau) + F_3(\tau)]V(\tau). \quad (18)$$

The solution of this equation can be obtained as

$$\begin{aligned} V(\tau) &= \exp \{ \Delta\tau [F_0(\tau_{n-1}) + F_1(\tau_{n-1}) + F_2(\tau_{n-1}) + F_3(\tau_{n-1})] \} V(\tau_{n-1}) \\ &= e^{\Delta\tau F_0(\tau_{n-1})} e^{\Delta\tau F_1(\tau_{n-1})} e^{\Delta\tau F_2(\tau_{n-1})} e^{\Delta\tau F_3(\tau_{n-1})} V(\tau_{n-1}) + O(\Delta\tau). \end{aligned} \quad (19)$$

Alternatively, a Páde approximation (1,0) can also be applied to all exponentials in Eq.(19) providing same order of approximation in $\Delta\tau$ but making all steps implicit. Namely, this results to the following splitting scheme of the solution of Eq.(18):

$$\begin{aligned} [1 - \Delta\tau F_0(\tau)]V^0 &= V(\tau_{n-1}), \\ [1 - \Delta\tau F_1(\tau)]V^1 &= V^0(\tau_{n-1}), \\ [1 - \Delta\tau F_2(\tau)]V^1 &= V^1(\tau_{n-1}), \\ [1 - \Delta\tau F_3(\tau)]V(\tau) &= V^2(\tau_{n-1}). \end{aligned} \quad (20)$$

We already know how to solve the first step in Eq.(20) (which always was a bottleneck for applying this fully implicit scheme). The remaining steps can be done similar to the steps 2-4 in the HV scheme. Thus, the whole first step in the HV scheme becomes implicit while has the same linear complexity in the number of nodes. Also our experiments confirm that this scheme provides great stability and preserves positivity of the solution. Therefore, running first few Rannacher steps is not necessary.

The third line in Eq.(7) can be modified accordingly as follows:

$$\begin{aligned} \tilde{Y}_0 &= Y_0 + \frac{1}{2}\Delta\tau [F(\tau_n)Y_k - F(\tau_{n-1})V_{n-1}], \\ &= Y_0 + \frac{1}{2}[Y_3 + \Delta\tau F(\tau_n)Y_3] - \frac{1}{2}[V_{n-1} + \Delta\tau F(\tau_{n-1})V_{n-1}] - \frac{1}{2}Y_3 + \frac{1}{2}V_{n-1} \\ &= Y_0 + \frac{1}{2}[\tilde{Y}_3 - Y_0 - Y_3 + V_{n-1}]. \end{aligned} \quad (21)$$

Here all values in the rhs of this equation are already known except of \tilde{Y}_3 which is the solution of the problem $\frac{\partial Y_3}{\partial \tau} = F(\tau_n)Y_3$. Therefore, it can be solved in the exactly same way as the first step of our fully implicit scheme.

Note that both Scheme A and Scheme B can be used as a part of the fully implicit scheme. If one makes a choice in favor of Scheme A, then the situation is as follows. We have the first step of the fully implicit scheme done with approximation $O(\sqrt{\Delta\tau} \min(h_S, h_v))$, and then multiple steps (six for the 3D problem) with the second order in space, so the total approximation is expected to be close to two. For the second sweep of the HV scheme this

is same. If the choice is made in favor of Scheme B, the rigorous spatial approximation of the whole HV scheme becomes two.

An obvious disadvantage of the proposed schemes is some degradation of performance, since it requires some number of iterations to converge⁶ when computing the mixed derivatives step, and at every iteration we need to solve 2 systems of linear equations. Despite the total complexity is still linear in the number of nodes, it takes about 4 times more computational time than the explicit scheme. However, as we have already mentioned, in our experiments the explicit scheme at the first and the third steps of Eq.(7) suffers from the curse of dimensionality, i.e., either it requires a very small and impractical temporal step to converge, or, otherwise, it explodes. Also our results show that the proposed scheme is only about 50-70% slower than the explicit step of the original HV scheme. However, the time step of our scheme can be significantly increased with no any stability issues, while this could be problematic for the HV scheme. Therefore, this increase in the time step could compensate the extra time required for doing the first step implicitly. For instance, running one step in time for the 3D advection-diffusion problem using the HV scheme coded in Matlab, at our machine takes 2 secs, while the fully implicit scheme requires 2.6 secs. On contrary, the HV scheme behaves kind of unstable with no Rannacher steps even with $\Delta\tau = 0.005$ yrs, while the fully implicit scheme continues to work well, e.g., with $\Delta\tau = 0.05$ yrs⁷. So if by the accuracy reason this step is sufficient, it can improve performance by factor 10, and then loosing about 50-70% for the implicit scheme is not sensitive.

Note, that the same approach can be applied to the forward equations based on the approach of [Itkin, 2015]. In more detail this will be presented elsewhere.

Once this step is accomplished, the whole scheme Eq.(7) becomes positivity-preserving. This is because for the first step of Eq.(7) our fully implicit scheme does preserve positivity. The next steps can be re-written in the form

$$[1 - \theta\Delta F_j(\tau_n)]Y_j = Y_{j-1} - \theta\Delta\tau F_j(\tau_{n-1})V_{n-1}$$

So $\Delta\tau$ can always be chosen small enough to make this step positivity-preserving, if the lhs matrix $M = 1 - \theta\Delta F_j(\tau_n)$ is an EM-matrix. The latter can be achieved by taking an approach of [Itkin, 2014a], where the first spatial derivatives are discretized by using one-sided finite-differences with the second order of approximation. Same is true for the second sweep of the splitting steps in Eq.(7).

2.2 Jump steps

Obtaining solutions at the 2-4 and 6-8 steps for some popular Lévy models such as Merton, Kou, CGMY (or GTSP), NIG, General Hyperbolic and Meixner ones, could be done as it is shown in [Itkin, 2014a, Itkin, 2014b]. Efficiency of this method in general is not worse than that of the Fast Fourier Transform (FFT), and in many cases is linear in N - the number of

⁶In our experiments 1-2 iterations were sufficient to provide the relative tolerance to be 10^{-6} , but the rate of this convergence should be studied in more detail. We leave it for the future investigation.

⁷Using the same β as in the above. However, changing the first multiplier in the rhs of Eq.(17) can make the scheme working for the higher values of the time step as well.

the grid points. In particular, this is the case for the Merton, Kou, CGMY/GTSP at $\alpha \leq 0$ and Meixner models.

Let sequentially consider all jump steps of the splitting algorithms.

2.2.1 Idiosyncratic jumps

Remember, that the characteristic exponent of the Meixner process is

$$\phi(u, a, b, d, m) = 2d \left\{ \log[\cos(b/2)] - \log \left[\cosh \left(\frac{au - ib}{2} \right) \right] \right\} + imu, \quad (22)$$

and the Lévy density $\mu(dy)$ of the Meixner process reads

$$\mu(dy) = d \frac{\exp(by/a)}{y \sinh(\pi y/a)} dy.$$

Therefore, from Eq.(5) we immediately obtain

$$\mathcal{J} = \phi(-i\nabla, a, b, d, m) = 2d \left\{ \log[\cos(b/2)] - \log \left[\cos \left(\frac{a\nabla + b}{2} \right) \right] \right\} + m\nabla. \quad (23)$$

The discretization scheme for this operator which provides second order of approximation in space and time while preserves positivity of the solution is given in [Itkin, 2014b], Propositions 3.8, 3.9.

At the end of this section we also remind, that according to the method of [Itkin, 2014b] the drift term in Eq.(23) (the last one) could be either moved into the drift term of the corresponding advection-diffusion operator, or could be discretized as

$$e^{\Delta\tau m\nabla_x} = \begin{cases} e^{\Delta\tau m A_{2,x}^F} + O(h_x^2), & m > 0, \\ e^{\Delta\tau m A_{2,x}^B} + O(h_x^2), & m < 0. \end{cases} \quad (24)$$

This is possible because in both cases in Eq.(24) the exponent is the negative of the EM-matrix⁸, therefore $e^{\Delta\tau m\nabla_x}$ is a positive matrix with all eigenvalues $|\lambda_i| < 1$.

2.2.2 Common jumps

The most difficult step is to solve the problem

$$V(\tau + \Delta\tau) = e^{\Delta\tau \mathcal{J}_{123}} V(\tau). \quad (25)$$

In [Itkin and Lipton, 2015] it was demonstrated how to do this when the common jumps are represented by the Kou model using a modification of the Peaceman-Rachford ADI method, see [McDonough, 2008]. Here we shortly describe the algorithm for the Meixner model.

⁸EM here stays for an eventually M-matrix, see [Itkin, 2014b].

Remember that by definition \mathcal{J}_{123} is given by Eq.(23) where now $\nabla = b_s \nabla_s + b_v \nabla_v + b_r \nabla_r$. The drift term $m \nabla$ again can be split among the corresponding drifts of the diffusion operators. After that we need to solve the following equation ([Itkin, 2014b])

$$\prod_{n=1}^{\infty} M_n^\kappa V(\tau + \Delta\tau) = [\cos(b/2)]^\kappa V(\tau), \quad (26)$$

$$M_n = 1 - \frac{(a \nabla + b)^2}{4\pi^2(n-1/2)^2}, \quad \kappa = 2d\Delta\tau,$$

where the parameters a, b, d characterize the common jumps.

This equation can be solved in a loop on n . Namely, we start with $n = 1$ and take $V_0 = [\cos(b/2)]^\kappa V(\tau)$. Since in our experiments $0 < \kappa < 1$,⁹ at every step in n we run this scheme for $\kappa = 0, 1$ and then use linear interpolation to κ . At $\kappa = 0$ an obvious solution is $V(\tau + \Delta\tau) = V(\tau)$. At $\kappa = 1$ Eq.(26) is a 3D parabolic equation that can be solved using our implicit version of the HV scheme. Indeed, it can be re-written in the form

$$\left[1 - (\Delta\tau)^2 K_n \left(\nabla + \frac{b}{a} \right)^2 \right] V = V(\tau), \quad K_n = \frac{a^2}{4\pi^2(n-1/2)^2(\Delta\tau)^2}.$$

As usually a is small, e.g., in [Schoutens, 2001] $a = 0.04$, so even for $n = 1$ $K_n = O(1)$. Now using the Páde approximation theory we can re-write this equation as

$$V = e^{(\Delta\tau)^2 K_n (\nabla + \frac{b}{a})^2} V(\tau) + O((\Delta\tau)^2).$$

Therefore, if we omit the last term $O((\Delta\tau)^2)$, the total second order approximation of the scheme in time is preserved. This latter equation is equivalent to

$$\frac{\partial V}{\partial s} = \left(\nabla + \frac{b}{a} \right)^2 V, \quad V(0) = \cos(b/2) V(\tau), \quad s \in [0, T_n], \quad (27)$$

which has to be solved at the time horizon (maturity) $T_n = (\Delta\tau)^2 K_n = \frac{a^2}{4\pi^2(n-1/2)^2}$. Since T_1 is small and usually less than $\Delta\tau$ we may solve it in one step in time. And when n increases, this conclusion remains to be true as well.

Once this solution is obtained we proceed to the next n . Thus, this scheme runs in a loop starting with $n = 1$ and ending at some $n = M$. Similar to how we did it for the idiosyncratic jumps we choose $M = 10$ based on the argument of [Itkin, 2014b], namely: i) the high order derivatives of the option price drop down pretty fast in value, and ii) first 10 terms of the sum $\sum_{i=1}^{\infty} T_i$ approximate the whole sum with the accuracy of 1%. The solution obtained after M steps is the final solution.

Overall, the whole splitting algorithm contains 11 steps. The complexity of each step is linear in N since at every step we solve some parabolic equation with a tridiagonal or pentadiagonal matrix. Thus, the total complexity of the method is $\varsigma N_1 N_2 N_3$ where N_i is

⁹This can always be achieved by choosing a relatively small $\Delta\tau$.

the number of grid nodes in the i -th dimension, and ς is some constant coefficient, which is about 276 (18 systems for one diffusion step if the implicit modification of the HV scheme is used times 2 diffusion steps, so totally 36; 10 systems per a 1D jump step times 2 steps times 3 variables, so totally 60; 18 steps per a single 3D parabolic PDE solution for common jumps times 10 steps, so totally 180).

Still this could be better than a straightforward application of the FFT (in case the FFT is applicable, e.g., the whole characteristic function is known in closed form which is not the case if one takes into account local volatility, etc.) which usually requires the number of FFT nodes to be a power of 2 with a typical value of 2^{11} . It is also better than the traditional approach which considers approximation of the linear non-local jump integral \mathcal{J} on some grid and then makes use of the FFT to compute a matrix-by-vector product. Indeed, when using FFT for this purpose we need two sweeps per dimension using a slightly extended grid (with, say, the tension coefficient ξ) to avoid wrap-around effects, [d’Halluin et al., 2005]. Therefore the total complexity per time step could be at least $O(8\xi_1\xi_2\xi_3N_1N_2N_3\log_2(\xi_1\xi_2\xi_3N_1N_2N_3))$ which for the FFT grid with $N_1 = N_2 = N_3 = 2048$, and $\xi_1 = \xi_2 = 1.3$ is 2.5 times slower than our method¹⁰. Also the traditional approach experiences some other problems for jumps with infinite activity and infinite variation, see survey in [Itkin, 2014a] and references therein. Also as we have already mentioned using Fast Gauss Transform for the common jump step could significantly reduce the time for this most time-consuming piece of the splitting scheme.

3 Numerical experiments

Due to the splitting nature of our entire algorithm represented by Eq.(6), each step of splitting is computed using a separate numerical scheme. All schemes provide second order approximation in both space and time, are unconditionally stable and preserve positivity of the solution.

In our numerical experiments for the steps which include mixed derivatives terms we used the suggested fully implicit version of the Hundsdorfer-Verwer scheme. This allows one to eliminate any additional damping scheme of the lower order of approximation, e.g., implicit Euler scheme (as this is done in the Rannacher method), or Do scheme with the parameter $\theta = 1$ (as this was suggested in [Haentjens and In’t Hout, 2012]).

A non-uniform finite-difference grid is constructed similar to [In’t Hout and Foulon, 2010] in v and r domains, and as described in [Itkin and Carr, 2011] in the S domain. In case of barrier options we extended the S grid by adding 2-3 ghost points either above the upper barrier or below the lower barrier, or both with the same boundary conditions as at the barrier (rebate or nothing). Construction of the jump grid, which is a superset of the finite-difference grid used at the first (diffusion) step is also described in detail in [Itkin, 2014a]. Normally the diffusion grid contained 61 nodes in each space direction. The extended jump grid contained extra 20-30 nodes. If a typical spot value at time $t = 0$ is $S_0=100$, the full

¹⁰As mentioned by the referee, since the flop counts rarely predict accurately an elapsed time, this statement should be further verified.

grid ended up at $S = 10^3$.

We computed our results in Matlab at a standard PC with Intel Xeon E5620 2.4 Ghz CPU. A typical elapsed time for computing one time step for the pure diffusion model with no jumps is given in the Table 1¹¹: Here $k = \log[t_i/t_{i-1}]/\log[N_i/N_{i-1}]$ is the power in

N of nodes	Advection-Diffusion			
	Mixed der	1D steps	Total for 1 sweep	k
50x50x50	0.81	0.38	1.19	-
60x60x60	1.26	0.59	1.85	2.42
70x70x70	1.88	0.86	2.74	2.54
80x80x80	2.71	1.28	3.89	2.62
100x100x100	4.50	2.22	3.89	3.17

Table 1: Elapsed time in secs for 1 step in time to compute the advection-diffusion problem.

the complexity C of calculations, which is regressed to $C \propto N^k$. It can be seen that the complexity is almost linear in all dimensions regardless of the number of nodes. The slight grows of k can be attributed to the way how Matlab processes large sparse matrices. Also our C++ implementation is about 15 times faster than Matlab.

European call option. In this test we solved an European call option pricing problem using the described model in a pure diffusion context, hence all jump intensities are set to zero. Also for simplicity we assumed all parameters of the model to be time-independent. Thus, in this test the robustness of our convection-diffusion FD scheme is validated¹². Parameters of the model used in this test are given in Table 2, and the results are presented in Fig. 1,2,3. We choose $a = b = 0.5$, $c = 1$, and the local volatility function was set to 1, so pretty much in this test our model is an lognormal + double CIR model (with stochastic volatility and interest rates).

T	K	κ_V	ξ_v	θ_v	κ_r	ξ_r	θ_r	q	ρ_{Sv}	ρ_{Sr}	ρ_{vr}	ϕ_{Sv}	α	β
1	100	2	0.3	0.9	3	0.1	0.05	0.5	-0.647	0	0.1	$4\pi/5$	0.5	0.5

Table 2: Parameters of the test for pricing an European call option.

We recall that a correlation matrix Σ of N assets can be represented as a Gram matrix with matrix elements $\Sigma_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ where $\mathbf{x}_i, \mathbf{x}_j$ are unit vectors on a $N - 1$ dimensional hyper-sphere \mathbb{S}^{N-1} . Using the 3D geometry, it is easy to establish the following cosine law for the correlations between three assets:

$$\rho_{xy} = \rho_{yz}\rho_{xz} + \sqrt{(1 - \rho_{yz}^2)(1 - \rho_{xz}^2)}\cos(\phi_{xy}),$$

¹¹Note, that, e.g., for the HV scheme we need 2 sweeps per one step in time.

¹²In this paper we don't analyze the convergence and order of approximation of the FD scheme, since the convergence in time is same as in the original HV scheme, and approximation was proven by the Theorem. For the jump FD schemes the convergence and approximation are considered in [Itkin, 2014a].

with ϕ_{xy} being an angle between \mathbf{x} and its projection on the plane spanned by \mathbf{y}, \mathbf{z} . As discussed, e.g., by [Dash, 2004], three variables $\rho_{yz}, \rho_{zz}, \phi_{xy}$ are independent, but $\rho_{xy}, \rho_{xz}, \rho_{yz}$ are not. Therefore, the value ρ_{Sv} in Table 2 was computed using given ρ_{Sr}, ρ_{vr} and ϕ_{Sv} .

Since the whole picture in this case is 4D, we represent it as a series of 3D projections, namely: Fig. 1 represents the $S - v$ plane at various values of the r coordinate which are indicated in the corresponding labels; Fig. 2 does same in the $S - r$ plane, and Fig. 3 - in the $v - r$ plane. Based on the values of parameters of the FD scheme the expected discretization error should be about 1%.

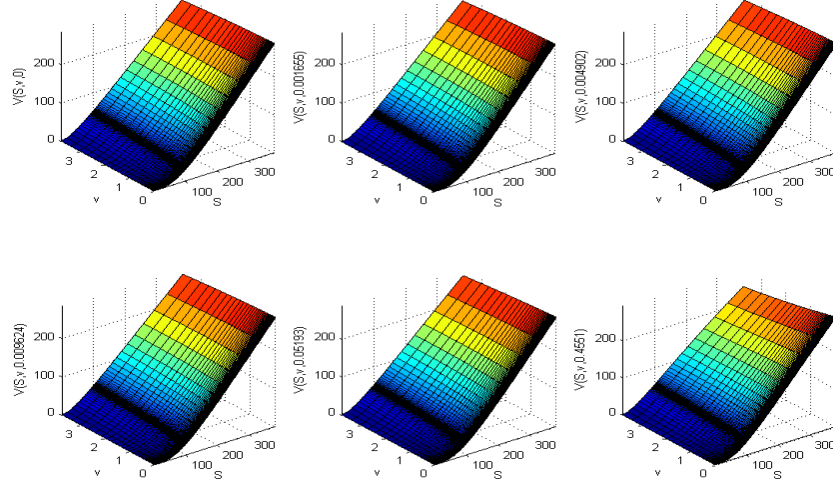


Figure 1: European call option prices in $S_0 - v_0$ plane at various values of r_0 .

Double barrier option. In this test we solved a more challenging problem of pricing double barrier option using the same model with no jumps with the lower barrier $L = 50$ and the upper barrier $H = 130$. Parameters of the model used in this test are given in Table 3, and the results are presented in Fig. 4,5,6.

T	K	κ_V	ξ_v	θ_v	κ_r	ξ_r	θ_r	q	ρ_{Sv}	ρ_{Sr}	ρ_{vr}	ϕ_{Sv}	α	β
0.5	100	2.5	0.3	0.6	0.3	0.1	0.05	0	-0.587	0.3	0.4	$4\pi/5$	0.5	0.5

Table 3: Parameters of the test for pricing a Double barrier call option.

It can be seen that the damping properties of the fully implicit HV scheme are sufficient, so no oscillations can be observed even near the critical points (close to strike and both barriers in S space)

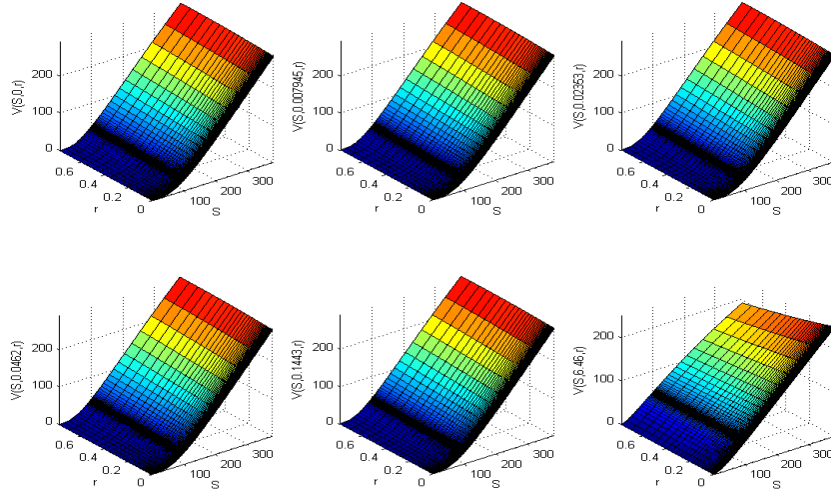


Figure 2: European call option prices in $S_0 - r_0$ plane at various values of v_0 .

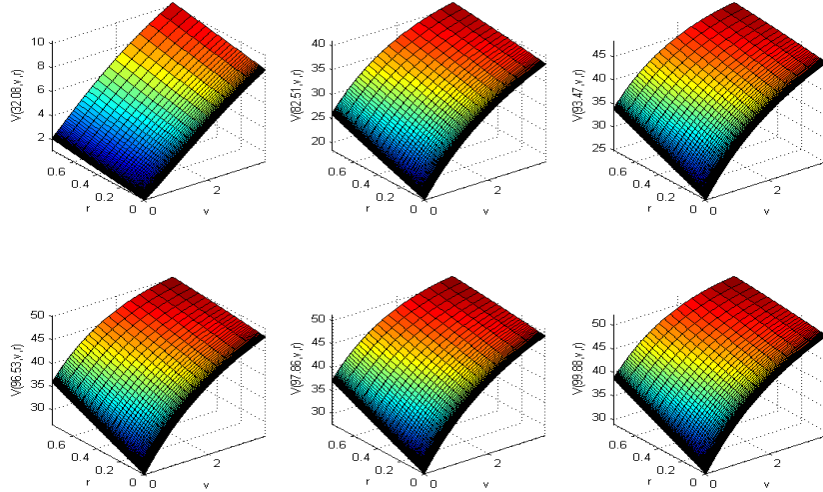


Figure 3: European call option prices in $v_0 - r_0$ plane at various values of S_0 .

Up-and-Out call option with jumps. The third test deals with jumps using the Meixner model for both idiosyncratic and common jumps as was described in the previous section. Parameters of the jump processes are given in Table 4, while the remaining parameters are the same as in Table 3. The loading factors (as they are defined in Proposition 1.1) we used in the test are: $b_S = 1$, $b_v = 2$, $b_r = 3$.

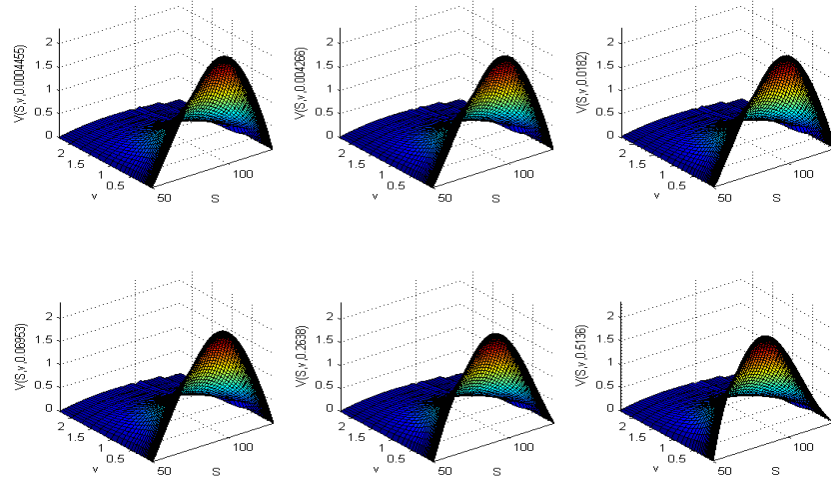


Figure 4: Double barrier call option prices in $S_0 - v_0$ plane at various values of r_0 .

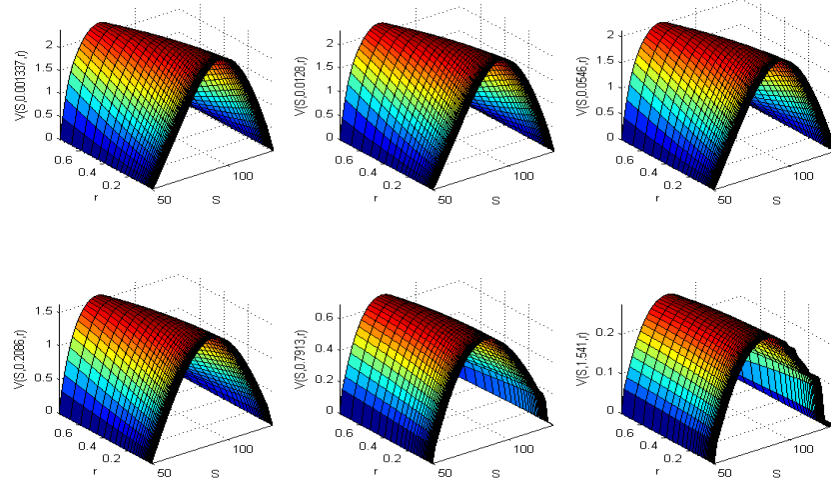


Figure 5: Double barrier call option prices in $S_0 - r_0$ plane at various values of v_0 .

A typical elapsed time for computing one time step for the pure jump model is given in the Table 5. Here we define the power k assuming that the complexity C is proportional to $(N_1 N_2 N_3)^\kappa$, so κ can be found as

$$\kappa = \log \left(\frac{t_i}{t_{i-1}} \right) / \log \left(\frac{N_{1i} N_{2i} N_{3i}}{N_{1,i-1} N_{2,i-1} N_{3,i-1}} \right).$$

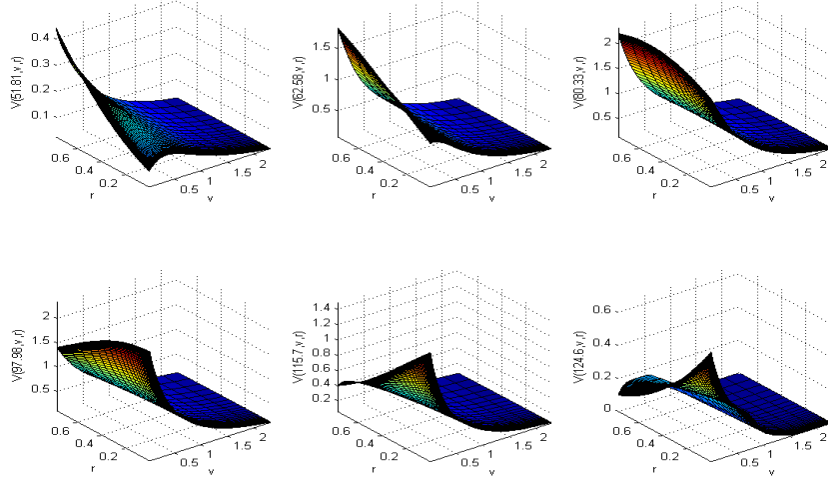


Figure 6: Double barrier call option prices in $v_0 - r_0$ plane at various values of S_0 .

Driver	a	b	m	d
S	0.04	-0.33	0.1	52
v	0.02	-0.5	0.03	40
r	0.01	-0.2	0.01	30
Common jumps	0.03	-0.1	0.05	40

Table 4: Parameters of the jump models.

One can see that in all experiments κ is close to 1, so the complexity is linear in the number of nodes.

The results computed in this experiment are presented in Fig. 7,8,9 as a difference between the full case with the correlated jumps and diffusion and that with no jumps. It is clear that jumps can play a significant role changing the whole 4D profile of the option price. Varying the loading factors one can change the correlations between jumps, and thus affect the price in a significant degree. For instance, increasing all the loading factors in this experiment by factor 10 results to the decrease of the Up-and-Out option price almost to few cents. Thus, the proposed model is very flexible. However, calibration of all the model parameters can be very time-consuming. Therefore, it is better to calibrate various pieces of the model separately, as this was discussed, e.g., in [Ballotta and Bonfiglioli, 2014]. Namely, the idiosyncratic jumps first can be calibrated separately to some marginal distributions using the appropriate instruments. Then the parameters of the common jumps can be calibrated to the option prices, while keeping parameters of the idiosyncratic jumps fixed.

N of nodes	Jumps			
	Common	all 1D	T_{1s}	κ
114x95x84	70.6	3.26	77.1	-
128x95x84	80.1	4.72	89.5	1.29
142x95x84	84.9	5.22	95.3	0.60
156x95x84	91.7	5.83	103.4	0.87
114x95x84	70.6	3.26	77.1	-
114x109x84	80.3	4.63	89.9	1.12
114x123x84	91.9	5.30	102.5	1.09
114x136x84	101.6	5.92	113.4	1.01
114x95x84	70.6	3.26	77.1	-
114x95x98	79.7	4.69	89.1	0.94
114x95x111	89.3	5.27	99.8	0.91
114x95x123	98.2	5.88	110.0	0.95

Table 5: Elapsed time in secs for 1 full time step in time T_{1s} to compute the 3D jump problem.

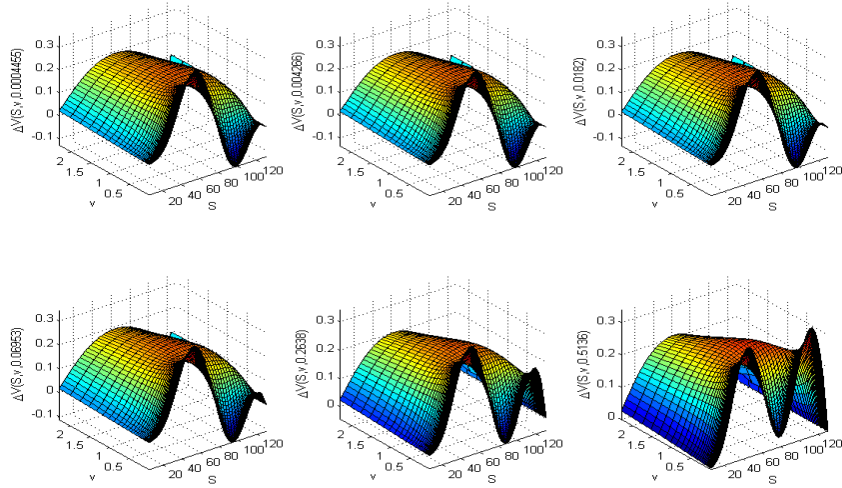


Figure 7: Difference in Up-and-Out call option prices computed with and without jumps in $S_0 - v_0$ plane at various values of r_0 .

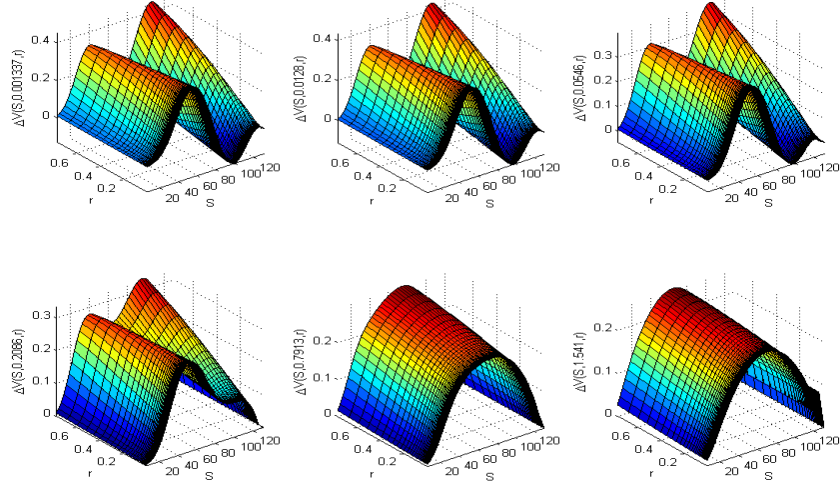


Figure 8: Difference in Up-and-Out call option prices computed with and without jumps in $S_0 - r_0$ plane at various values of v_0 .

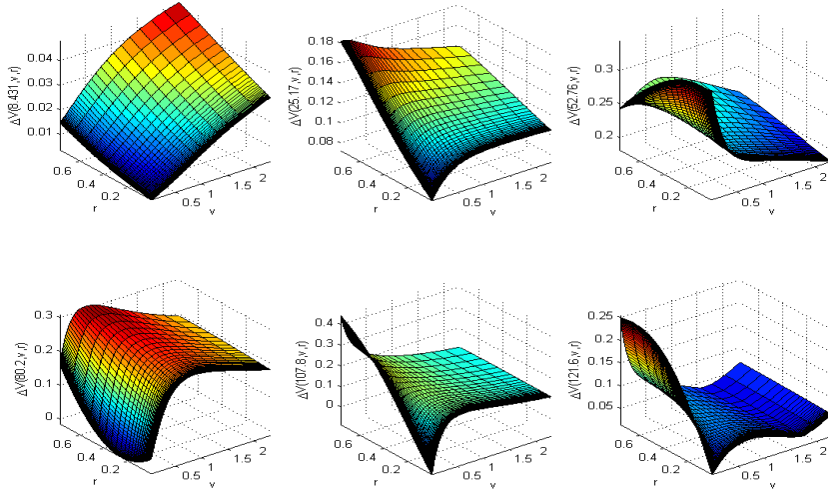


Figure 9: Difference in Up-and-Out call option prices computed with and without jumps in $v_0 - r_0$ plane at various values of S_0 .

4 Conclusion

In this paper we apply the approach of [Itkin and Lipton, 2015] for pricing credit derivatives to various option pricing problems (vanilla and exotic) where as an underlying model we use Local stochastic volatility model with stochastic interest rates and jumps in every stochastic driver. It is important that all jumps as well as the Brownian motions are correlated. Here we solve just the backward problem (solving the backward Kolmogorov equation, e.g., for pricing derivatives), while the forward problem (solving the forward Kolmogorov equation to find the density of the underlying process) can be treated in a similar way, see [Itkin, 2015].

In [Itkin and Lipton, 2015] test examples were given for the Kou and Merton models, while the approach is in no way limited by these models. Therefore, in this paper we demonstrate how a similar approach can be used together with the Meixner model. Again, this model is chosen only as an example, because, in general, the approach in use is rather universal. We provide an algorithm and results of numerical experiments.

The second contribution of the paper is a new fully implicit modification of the popular Hundsdorfer and Verwer and Modified Craig-Sneyd finite-difference schemes which provides second order approximation in space and time, is unconditionally stable and preserves positivity of the solution, while still keeps a linear complexity in the number of grid nodes. This scheme has extended damping properties, and, therefore, allows to eliminate any additional damping scheme of a lower order of approximation, e.g., implicit Euler scheme (as this is done in the Rannacher method), or Do scheme with the parameter $\theta = 1$ (as this was proposed in [Haentjens and In't Hout, 2012]). We prove unconditional stability of the scheme, second order of approximation in space and time and positivity of the solution. The results of our numerical experiments demonstrate the above conclusions.

To the best of author's knowledge both approaches have not been considered yet in the literature, so the main results of the paper are new.

The model in use is rather general, in a sense that it considers two (or even three) CEV processes for all the diffusion components and a wide class of the Lévy processes for the jump components. Therefore, a stable, accurate and sufficiently fast finite-difference approach for pricing derivatives using this model, which is proposed in this paper, could be beneficial for practitioners.

Acknowledgments

We thank Peter Carr, Alex Lipton and Alex Veygman for their useful comments and discussions. Also comments and suggestions of two anonymous referees are highly appreciated. We assume full responsibility for any remaining errors.

References

[Ballotta and Bonfiglioli, 2014] Ballotta, L. and Bonfiglioli, E. (2014). Multivariate asset models using Lévy processes and applications. *The European Journal of Finance*,

(DOI:10.1080/1351847X.2013.870917).

- [Bates, 1996] Bates, D. (1996). Jumps and stochastic volatility - exchange-rate processes implicit in deutschmark options. *Review of Financial Studies*, 9:69–107.
- [Boyarchenko and Levendorskii, 2013] Boyarchenko, S. and Levendorskii, S. (2013). American options in the Heston model with stochastic interest rate and its generalizations. *Applied Mathematical Finance*, 20(1):26–49.
- [Carr and Wu, 2004] Carr, P. and Wu, L. (2004). Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71:113–141.
- [Chen and Scott, 2004] Chen, R. and Scott, L. (2004). Stochastic volatility and jumps in interest rates: An international analysis. SSRN, 686985.
- [Chiarella and Kang, 2013] Chiarella, C. and Kang, B. (2013). The evaluation of American Compound option prices under stochastic volatility and stochastic interest rates. *Journal of Computational Finance*, 17(1):71–92.
- [Chiarella et al., 2008] Chiarella, C., Kang, B., Mayer, G., and Ziogas, A. (2008). The evaluation of American option prices under stochastic volatility and jump-diffusion dynamics using the method of lines. Technical Report Research paper 219, Quantitative Finance Research Centre, University of Technology, Sydney.
- [Cont and Tankov, 2004] Cont, R. and Tankov, P. (2004). *Financial modelling with jump processes*. Financial Mathematics Series, Chapman & Hall /CRCI.
- [Das, 2002] Das, S. R. (2002). The surprise element: jumps in interest rates. *Journal of Econometrics*, 106(1):27–65.
- [Dash, 2004] Dash, J. (2004). *Quantitative finance and risk management: a physicist's approach*. World Scientific.
- [d’Halluin et al., 2005] d’Halluin, Y., Forsyth, P. A., and Vetzal, K. R. (2005). Robust numerical methods for contingent claims under jump diffusion processes. *IMA J. Numerical Analysis*, 25:87–112.
- [Doffou and Hillard, 2001] Doffou, A. and Hillard, J. (2001). Pricing currency options under stochastic interest rates and jump-diffusion processes. *The Journal of Financial Research*, 25(4):565–585.
- [Durhama and Park, 2013] Durhama, G. and Park, Y. (2013). Beyond stochastic volatility and jumps in returns and volatility. *Journal of Business and Economic Statistics*, 31(1):107–121.
- [Gatheral, 2008] Gatheral, J. (2008). Consistent modeling of SPX and VIX options. In *Fifth World Congress of the Bachelier Finance Society*.

- [Giese, 2006] Giese, A. (2006). On the pricing of auto-callable equity structures in the presence of stochastic volatility and stochastic interest rates,. In *MathFinance Workshop*, Frankfurt. available at <http://www.mathfinance.com/workshop/2006/papers/giese/slides.pdf>.
- [Grzelak and Oosterlee, 2011] Grzelak, L. A. and Oosterlee, C. W. (2011). On the Heston model with stochastic interest rates,. *SIAM J. Finan. Math.*, 2:255–286.
- [Haentjens and In’t Hout, 2012] Haentjens, T. and In’t Hout, K. J. (2012). Alternating direction implicit finite difference schemes for the Heston–Hull–White partial differential equation. *Journal of Computational Finance*, 16:83–110.
- [Halperin and Itkin, 2013] Halperin, I. and Itkin, A. (2013). USLV: Unspanned Stochastic Local Volatility Model. available at <http://arxiv.org/abs/1301.4442>.
- [Hilpisch, 2011] Hilpisch, Y. (2011). Fast Monte Carlo valuation of American options under stochastic volatility and interest rates. available at http://www.google.com/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&cad=rja&uact=8&ved=OCB4QFjAA&url=http%3A%2F%2Farchive.euroscipy.org%2Ffile%2F4145%2Fraw%2FESP11-Fast_MonteCarlo_Paper.pdf&ei=8pR8VICxBsqbNpiYgvgB&usg=AFQjCNGWLwa5_cAsoSF0tJkRudUoD5V8jow.
- [Homescu, 2014] Homescu, C. (2014). Local stochastic volatility models: calibration and pricing. SSRN, 2448098.
- [Ikonen and Toivanen, 2007] Ikonen, S. and Toivanen, J. (2007). Componentwise splitting methods for pricing American options under stochastic volatility. *Int. J. Theor. Appl. Finance*, 10:331–361.
- [Ikonen and Toivanen, 2008] Ikonen, S. and Toivanen, J. (2008). Efficient numerical methods for pricing american options under stochastic volatility. *Num. Meth. PDEs*, 24:104–126.
- [In’t Hout and Foulon, 2010] In’t Hout, K. J. and Foulon, S. (2010). ADI finite difference schemes for option pricing in the Heston model with correlation. *International journal of numerical analysis and modeling*, 7(2):303–320.
- [In’t Hout and Welfert, 2007] In’t Hout, K. J. and Welfert, B. D. (2007). Stability of ADI schemes applied to convection-diffusion equations with mixed derivative terms. *Applied Numerical Mathematics*, 57:19–35.
- [Itkin, 2013] Itkin, A. (2013). New solvable stochastic volatility models for pricing volatility derivatives. *Review of Derivatives Research*, 16(2):111–134.
- [Itkin, 2014a] Itkin, A. (2014a). Efficient solution of backward jump-diffusion PIDEs with splitting and matrix exponentials. *Journal of Computational Finance*, Forthcoming. electronic version is available at <http://arxiv.org/abs/1304.3159>.

- [Itkin, 2014b] Itkin, A. (2014b). Splitting and matrix exponential approach for jump-diffusion models with Inverse Normal Gaussian, Hyperbolic and Meixner jumps. *Algorithmic Finance*, 3:233–250.
- [Itkin, 2015] Itkin, A. (2015). High-Order Splitting Methods for Forward PDEs and PIDEs. *International Journal of Theoretical and Applied Finance*, 18(5):1550031–1 —1550031–24.
- [Itkin and Carr, 2011] Itkin, A. and Carr, P. (2011). Jumps without tears: A new splitting technology for barrier options. *International Journal of Numerical Analysis and Modeling*, 8(4):667–704.
- [Itkin and Lipton, 2015] Itkin, A. and Lipton, A. (2015). Efficient solution of structural default models with correlated jumps and mutual obligations. *International Journal of Computer Mathematics*, DOI: 10.1080/00207160.2015.1071360.
- [Johannes, 2004] Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *The journal of Finance*, LIX(1):227–260.
- [Lipton, 2002] Lipton, A. (2002). The vol smile problem. *Risk*, pages 61–65.
- [McDonough, 2008] McDonough, J. M. (2008). *Lectures on computational numerical analysis of partial differential equations*. University of Kentucky. available at <http://www.engr.uky.edu/~acfd/me690-lctr-nts.pdf>.
- [Medvedev and Scaillet, 2010] Medvedev, A. and Scaillet, O. (2010). Pricing American options under stochastic volatility and stochastic interest rates. *Journal of Financial Economics*, 98:145–159.
- [Pagliarani and Pascucci, 2012] Pagliarani, S. and Pascucci, A. (2012). Approximation formulas for local stochastic volatility with jumps. SSRN, 2077394.
- [Rannacher, 1984] Rannacher, R. (1984). Finite element solution of diffusion equation with irregular data,. *Numerical mathematics*, 43:309–327.
- [Salmi et al., 2014] Salmi, S., Toivanen, J., and von Sydow, L. (2014). An IMEX-scheme for pricing options under stochastic volatility models with jumps. *SIAM J. Sci. Computing*, 36(4):B817–B834.
- [Schoutens, 2001] Schoutens, W. (2001). Meixner processes in finance. Technical report, K.U.Leuven–Eurandom.
- [Schoutens and Teugels, 1998] Schoutens, W. and Teugels, J. (1998). Lévy processes, polynomials and martingales. *Commun. Statist.- Stochastic Models*, 14(1,2):335–349.
- [Sepp, 2011a] Sepp, A. (2011a). Efficient numerical PDE methods to solve calibration and pricing problems in local stochastic volatility models. In *Global Derivatives*.

- [Sepp, 2011b] Sepp, A. (2011b). Parametric and non-parametric local volatility models: Achieving consistent modeling of VIX and equities derivatives. In *Quant Congress Europe*.
- [Sepp, 2014] Sepp, A. (2014). Log-normal stochastic volatility model: New insight and closed-form solution for vanilla options. Technical report, BAML.
- [Shirava and Takahashi, 2013] Shirava, K. and Takahashi, A. (2013). Pricing Basket options under local stochastic volatility with jumps. SSRN, 2372460.
- [Toivanen, 2010] Toivanen, J. (2010). A componentwise splitting method for pricing American options under the Bates model. In *Computational Methods in Applied Sciences*, pages 213–227. Springer.
- [Wade et al., 2005] Wade, B., A.Q.M.Khaliq, M.Siddique, and M.Yousuf (2005). Smoothing with positivity-preserving Páde schemes for parabolic problems with nonsmooth data. *Numerical Methods for Partial Differential Equations*, 21(3):553–573.

A Proof of Proposition 2.1

Recall that given the call option and positive correlation $\rho_{s,v} \geq 0$ we want to prove that the finite-difference scheme:

$$\begin{aligned} \left(QI_v + \sqrt{\Delta\tau}W(v)A_{2,v}^B \right) V^* &= \alpha^+ V(\tau) - V^k, \\ \left(PI_S - \sqrt{\Delta\tau}\rho_{s,v}W(S)A_{2,S}^F \right) V^{k+1} &= V^*. \\ \alpha^+ &= (PQ + 1)I_v - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_S^B + P\sqrt{\Delta\tau}W(v)A_v^F. \end{aligned} \quad (28)$$

is unconditionally stable in time step $\Delta\tau$, approximates Eq.(12) with $O(\sqrt{\Delta\tau} \max(h_S, h_v))$ and preserves positivity of the vector $V(x, \tau)$ if $Q = \beta\sqrt{\Delta\tau}/h_v$, $P = \beta\sqrt{\Delta\tau}/h_S$, where h_v, h_S are the grid space steps correspondingly in v and S directions, and the coefficient β must be chosen to obey the condition:

$$\beta > \max_{S,v} [W(v) + \rho_{s,v}W(S)].$$

First, let us show how to transform Eq.(11) to Eq.(28). Observe, that Eq.(11) can be re-written in the form

$$\begin{aligned} \left(P - \sqrt{\Delta\tau}\rho_{s,v}W(S)\nabla_S \right) \left(Q + \sqrt{\Delta\tau}W(v)\nabla_v \right) V^{(1)} &= V(\tau) - V^{(1)} + \alpha V^{(1)} \\ &= (\alpha + 1)V(\tau) - V^{(1)} + \alpha[V^{(1)} - V(\tau)], \\ \alpha &= PQ - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)\nabla_S + P\sqrt{\Delta\tau}W(v)\nabla_v. \end{aligned} \quad (29)$$

According to Eq.(9), $V^{(1)} - V(\tau) = \Delta\tau F_{Sv}(\tau) + O((\Delta\tau)^2)$. Also based on the proposition statement, $P \propto \sqrt{\Delta\tau}$, $Q \propto \sqrt{\Delta\tau}$, therefore $\alpha[V^{(1)} - V(\tau)] = O((\Delta\tau)^2)$. As we need just the first order approximation of Eq.(9), this term in Eq.(29) can be omitted. This gives rise to Eq.(28).

Non-negativity Now prove non-negativity of the solution. Consider first iteration ($k=0$), so the rhs of the first line in Eq.(28) can be written as $M_R V^0$, where

$$M_R \equiv PQ - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_S^B + P\sqrt{\Delta\tau}W(v)A_v^F.$$

Based on the definitions of the discrete operators A^F, A^B given right before the Proposition 2.1, one can see that matrix M_R has all non-negative elements outside of the main diagonal. The elements at the main diagonal $d_0(M_R)$ read

$$d_0(M_R) = PQ - \frac{3}{2}\sqrt{\Delta\tau} \left[\frac{Q\rho_{s,v}W(S)}{h_S} + \frac{PW(v)}{h_v} \right],$$

and are positive if

$$PQ > \frac{3}{2}\sqrt{\Delta\tau} \left[\frac{Q\rho_{s,v}W(S)}{h_S} + \frac{PW(v)}{h_v} \right].$$

This can be easily achieved if we put $Q = \beta\sqrt{\Delta\tau}/h_v$, $P = \beta\sqrt{\Delta\tau}/h_S$. The coefficient β must be chosen to obey the condition:

$$\beta > \frac{3}{2} \max_{S,v} [W(v) + \rho_{s,v}W(S)] > 0, \quad (30)$$

which guarantees that $d_0(M_R) > 0$. Since we require that in the proposition statement, the rhs of Eq.(28) is a non-negative vector.

To prove the non-negativity of the solution, consider first the second line in Eq.(28). We need to show that the matrix $M_R^S \equiv PI_S - \sqrt{\Delta\tau}\rho_{s,v}W(S)A_{2,S}^F$ is an EM-matrix, see Appendix A in [Itkin, 2014b]. This can be done similar to the proof of Lemma A.2 in [Itkin, 2014b], if one observes that the diagonal elements of M_R^S are positive, i.e.

$$d_{i,i}(M_R^S) = \frac{\sqrt{\Delta\tau}}{h_S} \left(\beta + \frac{3}{2}\rho_{s,v}W(S_i) \right) > 0, \quad i = 1, N_S. \quad (31)$$

Since M_R^S is an EM-matrix, its inverse is a non-negative matrix, therefore product of a non-negative matrix by a non-negative vector results in a non-negative vector. Therefore, the non-negativity of the solution is proved.

Convergence of iterations Since M_R^S is an EM-matrix, all its eigenvalues are non-negative. Also, since this is an upper triangular matrix, its eigenvalues are $d_{i,i}(M_R^S)$, $i = 1, N$, where N is the size of M_R^S . Also, by the properties of an EM-matrix, matrix $\| (M_R^S)^{-1} \|$ is non-negative, with the eigenvalues $\lambda_i = 1/d_{i,i}(M_R^S)$, $i \in [1, N]$.

Now due to Eq.(30) introduce a coefficient $\gamma > 1$ such that

$$\beta = \gamma \frac{3}{2} \max_{S,v} [W(v) + \rho_{s,v}W(S)].$$

From Eq.(31) we have

$$d_{i,i}(M_R^S) > \frac{3}{2}\rho_{s,v}W(S_i) \frac{\sqrt{\Delta\tau}}{h_S} (\gamma + 1).$$

Thus, it is always possible to provide the condition $d_{i,i}(M_R^S) > 1$ by an appropriate choice of γ . Accordingly, this gives rise to the condition $|\lambda_i| < 1, \forall i \in [1, N]$. The latter means that the spectral norm $\| (M_R^S)^{-1} \| < 1$, and, thus, the map $V^k \rightarrow V^{k+1}$ is contractual. This is the sufficient condition for the Picard iterations in Eq.(28) to converge. Unconditional stability follows. Other details about EM-matrices and necessary lemmas again can be found in [Itkin, 2014b].

For the first line of Eq.(28) we claim the same statement, i.e., that the matrix M_R^v is an EM-matrix. The main diagonal elements of M_R^v are also positive, namely

$$d_{j,j}(M_R^v) = \frac{\sqrt{\Delta\tau}}{h_v} \left(\beta + \frac{3}{2}W(v_j) \right) > 0, \quad j = 1, N_v.$$

The remaining proof again can be done based on definitions and Lemma A.2 in [Itkin, 2014b].

Thus, based on these two steps the coefficient γ has to be chosen to provide

$$\frac{3}{2} \max_{S,v} [W(v) + \rho_{s,v}W(S)] \frac{\sqrt{\Delta\tau}}{h_S} (\gamma + 1) > 1.$$

Since both steps on Eq.(28) converge in the spectral norm, and are unconditionally stable, the unconditional stability and convergence of the whole scheme follows. It also follows that the whole scheme preserves non-negativity of the solution.

Spatial approximation In Eq.(28) the lhs is approximated with the second order in h_S , while the first line in the rhs part uses the first order approximation of the first derivative. As $\nabla_S = A_S^B + O(h_S)$, and in the first line of the rhs of Eq.(28) we have a product $\sqrt{\Delta\tau}\nabla_S$, the order of the ignored terms is $O(\sqrt{\Delta\tau}h_S)$. So, rigorously speaking, the whole scheme Eq.(28) provides this order of approximation.

B Proof of Proposition 2.3

As compared with the Proposition 2.1, this scheme has the only modification. Namely, instead of the first step in Eq.(28)

$$\begin{aligned} (QI_v + \sqrt{\Delta\tau}W(v)A_{2,v}^B)V^* &= \alpha^+V(\tau) - V^k, \\ \alpha^+ &= (PQ + 1)I_v - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_S^B + P\sqrt{\Delta\tau}W(v)A_v^F. \end{aligned}$$

we now use

$$\begin{aligned} M^+ (QI_v + \sqrt{\Delta\tau}W(v)A_{2,v}^B)V^* &= V(\tau) - M^+V^k, \\ M^+ &= [\alpha^+]^{-1}, \quad \alpha^+ = (PQ + 1)I_v - Q\sqrt{\Delta\tau}\rho_{s,v}W(S)A_{2,S}^B + P\sqrt{\Delta\tau}W(v)A_{2,v}^F. \end{aligned} \tag{32}$$

This system can be solved by using the following fractional steps:

$$\begin{aligned}
\alpha^+ X &= V^k \\
Y &= V(\tau) - X, \\
\alpha^+ Z &= Y, \\
PV^* &= Z, \quad P \equiv \left(QI_v + \sqrt{\Delta\tau} W(v) A_{2,v}^B \right).
\end{aligned} \tag{33}$$

By construction, each matrix α^+ and P is an EM-matrix, see Appendix A in [Itkin, 2014a]. Therefore, steps 1,3,4 in Eq.(33) provide unconditional stability, and positivity of the solution. At step 2, the positivity can be provided by taking the time step $\Delta\tau$ sufficiently small. Also, the second order of approximation follows from the definition of $A_{2,S}^B$ and $A_{2,v}^F$. Moreover, in this case the approximation of α^+ is $O(\sqrt{\Delta\tau} \max(h_s^2, h_v^2))$, so the total approximation is determined by matrix P .

Since matrices α^+ and P are upper tridiagonal matrices, the complexity of solving Eq.(33) is linear in each direction.

The whole FD scheme in the Proposition 2.3 in addition to Eq.(32) also includes the second step which is same as in the Proposition 2.1. Since this step has the second order approximation in spatial variables, the whole scheme also provides the second order. The prove of convergence of the whole scheme is similar to that in Proposition 2.1.